

10 NO. 4135
ASTIA FILE COPY

AN EXTENSION OF MASSEY'S DISTRIBUTION OF THE MAXIMUM DEVIATION
BETWEEN TWO SAMPLE CUMULATIVE STEP FUNCTIONS

CHIA KUEI TSAO

TECHNICAL REPORT NO. 7

July 15, 1953

PREPARED UNDER CONTRACT NONR-491(00)
(NR-042-017)

FOR

OFFICE OF NAVAL RESEARCH

DEPARTMENT OF MATHEMATICS
WAYNE UNIVERSITY
DETROIT, MICHIGAN

AN EXTENSION OF MASSEY'S DISTRIBUTION OF THE MAXIMUM DEVIATION BETWEEN TWO SAMPLE CUMULATIVE STEP FUNCTIONS

By Chia Kuei Tsao
Wayne University

1. Summary and Introduction

Let $x_1 < x_2 < \dots < x_n$ and $y_1 < y_2 < \dots < y_m$ be the ordered results of two random samples from populations having continuous cumulative distribution functions $F(x)$ and $G(x)$ respectively. Let $S_n(x) = k/n$ where k is the number of observations of X which are less than or equal to x and $S'_m(x) = j/m$ where j is the number of observations of Y which are less than or equal to x .

The statistics

$$d_r = \max_{x \leq x_r} |S_n(x) - S'_m(x)|$$

and

$$d'_r = \max_{\substack{x \leq \max(x_r, y_r) \\ r \leq \min(m, n)}} |S_n(x) - S'_m(x)|$$

can be used to test the hypothesis $F(x) = G(x)$. For example, using d_r we would reject the hypothesis if the observed d_r , i.e., the maximum absolute deviation between the two step functions below the r th observation of a given sample, is significantly large.

2. Distribution of d_r

Denote by m_1 the number of observed values of Y which are less than x_1 , by m_2 the number of values of Y which are between x_1 and x_2 , ..., by m_r the number of values of Y which are between x_{r-1} and x_r and by M the number of values of Y which are greater than x_r . If the hypothesis $F(x) \equiv G(x)$ is true, the probability of the occurrence of a set of m_1, m_2, \dots, m_r, M can be shown to be

$$\Pr(m_1, \dots, m_r, M) = \frac{\binom{M+n-r}{M}}{\binom{m+n}{m}}$$

which depends only on M , i.e., for any given M the probability of the occurrence of any set of m_1, m_2, \dots, m_r is always $\binom{M+n-r}{M} / \binom{m+n}{m}$. Thus, for any given M , the probability that $d_r \leq a$ can be found by counting the number of sets of m_1, m_2, \dots, m_r which give values of $d_r \leq a$. Denote this number of sets by $K_{r,M}(a)$, then

$$\Pr(d_r \leq a) = \sum_{M=0}^m K_{r,M}(a) \frac{\binom{M+n-r}{M}}{\binom{m+n}{m}}.$$

The method of counting $K_{r,M}(a)$ is essentially the same as in [§5]. As an illustration, suppose $m=n$, then $S_m(x)$ and $S'_m(x)$ can only differ by multiples of $\frac{1}{m}$. For any integer c and any given M , we count the number of sets of m_1, m_2, \dots, m_r such that $d_r \leq \frac{c}{m}$.

Denote by $V_{ij}(c)$, $i = 1, 2, \dots, r$, $j = 1, 2, \dots, 2c$, the number of sets of possible m_1, m_2, \dots, m_i such that $d_i \leq \frac{c}{m}$. Then it is evident

that these $V_{ij}(c)$ satisfy the following difference equations

$$V_{ik}(c) = \sum_{j=1}^{k+1} V_{i-1,j}(c) \quad \begin{array}{l} i = 1, 2, \dots, r \\ k = 1, 2, \dots, 2c \end{array}$$

where

$$\begin{array}{ll} V_{0k}(c) = 1 & k \geq c, k = 1, 2, \dots, 2c \\ V_{0j}(c) = 0 & \text{otherwise.} \end{array}$$

Hence,

$$K_{r,M}\left(\frac{c}{M}\right) = V_{r, M-M-r+c+1}(c).$$

3. Distribution of d_r^1

If $r \leq n$ d_r , then a test based on the number of observations of one sample which are less than or equal to the r th observation of another sample becomes a one-sided test. To avoid this, we may use the statistic d_r^1 . In this case

$$\begin{aligned} \Pr(d_r^1 \leq a) &= \Pr(d_r \leq a, x_r > y_r) + \Pr(d_r \leq a, x_r < y_r) \\ &= \sum_{M=0}^{m-r} K_{r,M}(a) \frac{\binom{M+n-r}{M}}{\binom{m+n}{m}} \\ &\quad + \sum_{N=0}^{n-r} K_{r,N}(a) \frac{\binom{N+m-r}{N}}{\binom{m+n}{m}} \end{aligned}$$

where n_1, n_2, \dots, n_r, N have the same meaning as m_1, m_2, \dots, m_r, M .

If $m = n$, then

$$\Pr(d'_r \leq a) = 2 \sum_{M=0}^{m-r} K_{r,M}(a) \frac{\binom{M+m-r}{M}}{\binom{2m}{m}}.$$

We note that:

(a) If $r = m = n$, then both distributions of d_m and d'_m reduce to Massey's distribution [5].

(b) If $r = 1$, then d_r reduces to a special case of Gumbel and Schelling's exceedances problems [2].

Table I and II give the probabilities of d_r and d'_r respectively for $m = n$.

4. Applications

The statistics d_r and d'_r are useful for situations where the sample sizes are known, but the information beyond a certain ordered observation, say x_r , is unavailable. In life testing, one often wishes, by drawing two samples, to detect whether the mean life in one population is larger than that of another. If the observations become available in order of magnitude, then we can stop the experiment whenever at least r observations of each sample have occurred and reach a decision by the use of d'_r . Evidently, by doing so, it would be possible, in many cases, to reduce both the average time needed and/or the average number of items destroyed.

As an illustration, we give a numerical example as follows:

Suppose fuses are produced by two different methods. One is interested in detecting whether the mean current needed to blow the fuses produced by the

first method is different from that produced by the second. This can be considered as testing whether two populations are the same. To this end, one then may put, say, 40 fuses produced by the first method and another 40 by the second on a test. Suppose one arranges the test in such a way that every fuse in the 2 samples is subjected to the same current so that the weakest blows first, then the second, etc. Let us choose in advance that $r = 6$ and $\alpha = .05$. Let $x_1 < x_2 < \dots$ denote the ordered observed current needed to blow the fuses in the first sample and $y_1 < y_2 < \dots$ those in the second. Suppose that the actual combined outcomes are as follows: $x_1 x_2 x_3 x_4 y_1 x_5 x_6 x_7 y_2 x_8 x_9 y_3 x_{10} x_{11} x_{12} \dots$ then the experiment may be terminated when the observation x_{12} has occurred and reject the null hypothesis using the statistic d_6^1 , since for $m = n = 40$, $\Pr(d_6^1 \geq 9) = .0451$ from Table II. In this particular experiment, only 20% of the fuses are destroyed in reaching a decision.

It, perhaps, should be remarked that if we define

$$D_r = \max_{x \geq x_{n-r+1}} |S_n(x) - S_m^1(x)|$$

$$D_r^1 = \max_{\substack{x \geq (x_{n-r+1}, y_{m-r+1}) \\ r \leq \min(m, n)}} |S_n(x) - S_m^1(x)|$$

then the distribution of D_r , and D_r^1 are identical with those of d_r and d_r^1 . Thus, in a test, if the information below a certain ordered observation is unavailable, or if the observations become available in this order:

X_n, X_{n-1}, \dots, X_1 and Y_m, Y_{m-1}, \dots, Y_1 , then D_r or D_r' would be the appropriate statistic to use.

In conclusion, I would like to thank Mrs. Dorothy Wolfe who carried out the computations of Tables I and II.

References

1. W. Feller, "On the Kolmogorov-Smirnov Limit Theorems for Empirical Distributions", Annals of Math. Stat., Vol. 19, (1948), pp. 177-189.
2. E. J. Gumbel and H. von Schelling, "The Distribution of the number of exceedances", Annals of Math. Stat., Vol. 21 (1950), pp. 247-262.
3. F. J. Massey, Jr., "A note on the estimation of a distribution function by confidence limits", Annals of Math. Stat., Vol. 21, (1950), pp. 114-119.
4. F. J. Massey, Jr., "A note on the power of a non-parametric test", Annals of Math. Stat., Vol. 21, (1950), pp. 440-443.
5. F. J. Massey, Jr., "The distribution of the maximum deviation between two sample cumulative step functions", Annals of Math. Stat., Vol. 22, (1951), pp. 125-128.
6. N. Smirnov, "Tables for estimating the goodness of fit of empirical distributions", Annals of Math. Stat., Vol. 19, (1948), pp. 279-281.
7. S. S. Wilks, "Statistical prediction with special reference to the problem of tolerance limits", Annals of Math. Stat., Vol. 13, (1942), pp. 400-409.

Table I

Probability of $d_r \leq c/m$

m	r	1	2	3	4	5	6	7	8	9	10	11	12
3	2	.50000	.95000	1.00000									
4	2	.45714	.81429	.98571	1.00000								
	3	.28571	.81429	.98571	1.00000								
5	2	.43651	.85714	.96032	.99603	1.00000							
	3	.25397	.73810	.96032	.99603	1.00000							
	4	.15873	.67857	.93651	.99603	1.00000							
6	2	.42424	.84091	.94372	.98701	.99892	1.00000						
	3	.23810	.70130	.92857	.98701	.99892	1.00000						
	4	.13853	.60390	.89827	.98701	.99892	1.00000						
	5	.08658	.55519	.87229	.97944	.99892	1.00000						
7	2	.41608	.83042	.93240	.97931	.99592	.99971	1.00000					
	3	.22844	.67920	.90793	.97348	.99592	.99971	1.00000					
	4	.12821	.56643	.85897	.97348	.99592	.99971	1.00000					
	5	.07459	.53854 41776	.82634	.96300	.99592	.99971	1.00000					
	6	.04662	.44843	.80186	.95280	.99359	.99971	1.00000					

Table I (continued)

r	c											
	1	2	3	4	5	6	7	8	9	10	11	12
8	2	.41026	.82308	.92424	.97319	.99277	.99876	.99992	1.00000			
	3	.22191	.66434	.89378	.96232	.99068	.99876	.99992	1.00000			
	4	.12183	.54336	.83294	.95649	.99068	.99876	.99992	1.00000			
	5	.06838	.45315	.78291	.94367	.99068	.99876	.99992	1.00000			
	6	.03978	.39021	.75245	.92968	.98718	.99876	.99992	1.00000			
	7	.02486	.35874	.73007	.91880	.98345	.99806	.99992	1.00000			
9	2	.40588	.81765	.91810	.96833	.98992	.99757	.99963	.99998	1.00000		
	3	.21719	.65362	.88355	.95352	.98548	.99685	.99963	.99998	1.00000		
	4	.11748	.52756	.81473	.94272	.98322	.99685	.99963	.99998	1.00000		
	5	.06450	.43149	.75376	.92236	.98322	.99685	.99963	.99998	1.00000		
	6	.03620	.35985	.70769	.90539	.97869	.99685	.99963	.99998	1.00000		
	7	.02106	.30987	.68005	.89058	.97314	.99572	.99963	.99998	1.00000		
	8	.01316	.28488	.65981	.87982	.96870	.99443	.99942	.99998	1.00000		

Table I (continued)

m	r	1	2	3	4	5	6	7	8	9	10	11	12
10	2	.40248	.81347	.91331	.96440	.98744	.99637	.99921	.99989	.99999	1.00000		
	3	.21362	.64551	.87580	.94613	.98086	.99466	.99897	.99989	.99999	1.00000		
	4	.11431	.51602	.80128	.93192	.97610	.99383	.99897	.99989	.99999	1.00000		
	5	.06183	.41650	.73309	.90525	.97378	.99383	.99897	.99989	.99999	1.00000		
	6	.03395	.34065	.67739	.88049	.96836	.99383	.99897	.99989	.99999	1.00000		
	7	.01952	.28409	.63587	.86262	.96121	.99228	.99897	.99989	.99999	1.00000		
	8	.01108	.24464	.61101	.84804	.95164	.99020	.99861	.99989	.99999	1.00000		
	9	.00693	.22491	.59483	.83759	.94987	.98849	.99818	.99983	.99999	1.00000		
15	2	.39272	.80172	.89962	.95259	.97920	.99159	.99690	.99898	.99990	.99993	.99999	1.00000
	3	.20383	.62328	.85472	.92635	.96571	.98547	.99447	.99815	.99947	.99987	.99998	1.00000
	4	.10611	.48591	.76593	.90199	.95254	.97933	.99203	.99735	.99926	.99994	.99997	1.00000
	5	.05544	.38006	.68171	.85097	.94046	.97372	.98990	.99671	.99913	.99982	.99997	1.00000
	6	.02909	.29843	.60791	.81327	.92114	.96902	.98828	.99632	.99908	.99982	.99997	1.00000
	7	.01534	.23544	.51425	.77033	.90032	.96232	.98728	.99617	.99908	.99982	.99997	1.00000
	8	.00814	.18683	.48977	.73200	.88009	.95482	.98578	.99617	.99908	.99982	.99997	1.00000
	9	.00436	.14935	.44362	.69904	.86242	.94774	.98378	.99582	.99908	.99982	.99997	1.00000
10	10	.00236	.12055	.40521	.67192	.84825	.94159	.98135	.99515	.99898	.99982	.99997	1.00000

Table I (continued)

	1	2	3	4	5	6	7	8	9	10	11	12
1	.38808	.79626	.89313	.94674	.97476	.98868	.99519	.99808	.99928	.99975	.99992	.99998
2	.19939	.61316	.84525	.91681	.95784	.97989	.99100	.99624	.99854	.99948	.99983	.99995
3	.10260	.47286	.75053	.88858	.94070	.97047	.98627	.99410	.99767	.99916	.99973	.99992
4	.05289	.36526	.66048	.83900	.92417	.96103	.98143	.99186	.99674	.99883	.99963	.99990
5	.02733	.28268	.58122	.78559	.89854	.95116	.97672	.98969	.99580	.99853	.99954	.99988
6	.01414	.21923	.51231	.73418	.86907	.93937	.97236	.98773	.99512	.99829	.99948	.99987
7	.00734	.17043	.45256	.68657	.83906	.92458	.96661	.98607	.99153	.99813	.99945	.99986
8	.00382	.13287	.40080	.64319	.81020	.90909	.96046	.98417	.99114	.99804	.99944	.99986
9	.00200	.10393	.35605	.60405	.78336	.89401	.95397	.98198	.99011	.99802	.99944	.99986
10	.38359	.79103	.88687	.94096	.97024	.98549	.99316	.99688	.99863	.99942	.99976	.99991
11	.19520	.60360	.83640	.90768	.94998	.97393	.98693	.99370	.99708	.99870	.99944	.99977
12	.09940	.46086	.73634	.87616	.92922	.96122	.97963	.98974	.99504	.99771	.99899	.99957
13	.05065	.35211	.64144	.82088	.90888	.94809	.97170	.98525	.99265	.99651	.99842	.99932
14	.02583	.26922	.55808	.76114	.87750	.93498	.96350	.98045	.99001	.99515	.99776	.99902
15	.01318	.20600	.48576	.70332	.84110	.91690	.95526	.97550	.98724	.99369	.99705	.99870
16	.00673	.15775	.42312	.64937	.80340	.89530	.94486	.97055	.98442	.99220	.99631	.99836
17	.00344	.12092	.34041	.59969	.76628	.87193	.93244	.96472	.98164	.99072	.99559	.99803
18	.00176	.09278	.32189	.55418	.73062	.84800	.91867	.94868	.97355	.98930	.99490	.99773

Table I (continued)

	2	3	4	5	6	7	8	9	10	11	12
1 .38039	.78851	.88382	.93811	.96793	.98361	.99203	.99617	.99821	.99918	.99963	.99994
2 .00319	.59901	.83218	.90327	.94607	.97086	.98472	.99221	.99614	.99814	.99913	.99960
3 .09790	.45521	.72965	.87031	.92365	.95655	.97606	.98722	.99339	.99668	.99838	.99924
4 .04962	.34605	.72363	.81244	.90164	.94169	.96659	.98148	.99007	.99484	.99741	.99874
5 .02517	.25940	.54759	.74992	.86770	.92678	.95669	.97524	.98633	.99270	.99623	.99812
6 .01277	.20022	.47403	.68948	.82829	.90623	.94661	.96869	.98228	.99031	.99489	.99740
7 .00648	.15239	.41049	.63311	.78741	.88163	.93389	.96198	.97803	.98776	.99342	.99660
8 .00329	.11604	.35572	.58144	.74735	.85523	.91899	.95439	.97402	.98543	.99221	.99602
9 .00167	.08840	.30821	.53369	.70817	.82735	.90161	.94461	.96875	.98234	.99025	.99482

Table II

Probability of $d_p' \leq c/m$

c	1	2	3	4	5	6	7	8	9	10	11	12
0	1											
1												
2	.140000	.900000	1.000000									
3												
4	.314286	.77143	.97143	1.000000								
5	.52657	.77143	.97143	1.000000								
6												
7	.317146	.71429	.92063	.99206	1.000000							
8	.19048	.64286	.92063	.99206	1.000000							
9	.12696	.64286	.92063	.99206	1.000000							
10												
11	.230303	.68182	.88745	.97103	.99784	1.000000						
12	.17316	.58442	.85714	.97103	.99784	1.000000						
13	.10390	.52597	.85714	.97103	.99784	1.000000						
14	.06926	.52597	.85714	.97103	.99784	1.000000						
15												
16	.29371	.66084	.86480	.95862	.99184	.99942	1.000000					
17	.16317	.55070	.81585	.94697	.99184	.99942	1.000000					
18	.09324	.47203	.78788	.94697	.99184	.99942	1.000000					
19	.05594	.42483	.78788	.94697	.99184	.99942	1.000000					
20	.03730	.42483	.78788	.94697	.99184	.99942	1.000000					

Table II (continued)

	1	2	3	4	5	6	7	8	9	10	11	12
2	.28718	.64615	.84848	.91639	.98555	.99751	.99984	1.00000				
3	.15664	.52867	.76717	.92463	.98135	.99751	.99984	1.00000				
4	.08702	.44056	.74281	.91298	.98135	.99751	.99984	1.00000				
5	.04973	.37762	.71733	.91298	.98135	.99751	.99984	1.00000				
6	.02828	.33986	.71733	.91298	.98135	.99751	.99984	1.00000				
7	.01989	.33986	.71733	.91298	.98135	.99751	.99984	1.00000				
2	.28235	.63529	.83620	.93665	.97984	.99515	.99926	.99996	1.00000			
3	.15204	.51312	.76709	.90703	.97096	.99371	.99926	.99996	1.00000			
4	.08293	.41983	.71181	.88544	.96643	.99371	.99926	.99996	1.00000			
5	.04607	.34986	.67133	.87413	.96643	.99371	.99926	.99996	1.00000			
6	.02633	.29988	.64829	.87413	.96643	.99371	.99926	.99996	1.00000			
7	.01644	.26989	.64829	.87413	.96643	.99371	.99926	.99996	1.00000			
8	.01053	.26989	.64829	.87413	.96643	.99371	.99926	.99996	1.00000			

Table II (continued)

	1	2	3	4	5	6	7	8	9	10	11	12
10	.27864	.62594	.82663	.92879	.97487	.99274	.99842	.99978	.99999	1.00000		
9	.14861	.50155	.75161	.89307	.96172	.98933	.99794	.99978	.99999	1.00000		
8	.08002	.40510	.68926	.86384	.95220	.98766	.99794	.99978	.99999	1.00000		
7	.04365	.31144	.63955	.84300	.94755	.98766	.99794	.99978	.99999	1.00000		
6	.02425	.27617	.60317	.83218	.94755	.98766	.99794	.99978	.99999	1.00000		
5	.01386	.23674	.58248	.83218	.94755	.98766	.99794	.99978	.99999	1.00000		
4	.00831	.21307	.58248	.83218	.94755	.98766	.99794	.99978	.99999	1.00000		
3	.00554	.21307	.58248	.83218	.94755	.98766	.99794	.99978	.99999	1.00000		
2	.26820	.60345	.79923	.90517	.95840	.98318	.99380	.99795	.99911	.99985	.99997	1.00000
1	.13946	.47069	.70945	.85270	.93142	.97094	.98894	.99629	.99894	.99975	.99995	.99999
	.07276	.36837	.63148	.80397	.90509	.95865	.98406	.99463	.99853	.99968	.99995	.99999
	.03811	.28943	.56419	.76009	.88099	.94744	.97979	.99341	.99826	.99965	.99995	.99999
	.02006	.22850	.50617	.72137	.85979	.93805	.97655	.99264	.99816	.99965	.99995	.99999
	.01062	.18145	.45702	.68715	.84224	.93096	.97455	.99234	.99816	.99965	.99995	.99999
	.00566	.14516	.41515	.66003	.82875	.92646	.97375	.99234	.99816	.99965	.99995	.99999
	.00305	.11725	.38017	.63798	.81978	.92454	.97375	.99234	.99816	.99965	.99995	.99999
10	.00166	.09593	.35340	.62247	.81558	.92454	.97375	.99234	.99816	.99965	.99995	.99999

Table II (continued)

	1	2	3	4	5	6	7	8	9	10	11	12
2	.26334	.59252	.78631	.89348	.94956	.97735	.99038	.99616	.99856	.99950	.99984	.99995
3	.13543	.45708	.69050	.83363	.91568	.95978	.98201	.99248	.99709	.99896	.99966	.99990
4	.06977	.35320	.60709	.77715	.88141	.94093	.97254	.98820	.99533	.99832	.99946	.99985
5	.03601	.27345	.53471	.72513	.84835	.92205	.96287	.98372	.99349	.99766	.99925	.99979
6	.01863	.21216	.47199	.67771	.81727	.90393	.95344	.97938	.99161	.99706	.99908	.99976
7	.00966	.16501	.41753	.63473	.78857	.88706	.94472	.97546	.99024	.99658	.99897	.99974
8	.00502	.12871	.37041	.59598	.76243	.87180	.93701	.97215	.98907	.99626	.99890	.99973
9	.00262	.10073	.32968	.56125	.73899	.85841	.93054	.96958	.98828	.99609	.99888	.99973
10	.00137	.07914	.29454	.53039	.71837	.84708	.92546	.96781	.98065	.99603	.99888	.99973
2	.25870	.58207	.77374	.88192	.94047	.97098	.98632	.99376	.99725	.99883	.99952	.99981
3	.13170	.44449	.67279	.81536	.89995	.94787	.97386	.98739	.99415	.99739	.99889	.99954
4	.06709	.33966	.58509	.75233	.85844	.92244	.95925	.97947	.99009	.99542	.99798	.99915
5	.03420	.25974	.50914	.69395	.81775	.89618	.94341	.97050	.98530	.99301	.99684	.99864
6	.01745	.19878	.44338	.64035	.77873	.86996	.92701	.96089	.98003	.99030	.99552	.99805
7	.00891	.15226	.38644	.59128	.74172	.84434	.91053	.95101	.97448	.98739	.99410	.99740
8	.00455	.11673	.33711	.54641	.70685	.81962	.89430	.94110	.96885	.98440	.99263	.99673
9	.00233	.08958	.29438	.50543	.67412	.79599	.87856	.93138	.96328	.98144	.99118	.99607
10	.00119	.06882	.25734	.46801	.64349	.76899	.86348	.90360	.95789	.97859	.98979	.99546

Table II (continued)

	1	2	3	4	5	6	7	8	9	10	11	12
1	.25645	.57702	.76765	.87621	.93586	.96763	.98407	.99235	.99641	.99836	.99927	.99968
2	.17994	.43853	.66436	.80654	.89214	.94172	.96943	.98442	.99228	.99628	.99826	.99921
3	.06586	.33341	.57487	.74061	.84730	.91310	.95212	.97444	.98677	.99336	.99577	.99848
4	.03339	.25358	.49756	.67961	.80328	.88338	.93318	.96297	.98014	.98969	.99482	.99748
5	.01694	.19294	.43082	.62363	.76095	.85356	.91337	.95049	.97265	.98540	.99247	.99625
6	.00860	.14626	.37319	.57241	.72071	.82419	.89321	.93739	.96455	.98063	.98978	.99480
7	.00436	.11184	.32341	.52560	.68267	.79564	.87308	.92396	.95606	.97551	.98685	.99320
8	.00222	.08521	.28061	.48327	.64740	.76874	.85391	.91114	.94804	.97086	.98443	.99216
9	.00113	.05496	.24324	.44375	.61307	.74161	.83380	.89688	.93852	.96468	.98051	.98964